



## SEQUENCE LISTING

&lt;110&gt; Strathmann AG &amp; Co.

&lt;120&gt; Virus-Vaccine

&lt;130&gt; P057760

&lt;140&gt;

&lt;141&gt;

&lt;150&gt; 199 07 485.2

&lt;151&gt; 1999-02-12

&lt;160&gt; 27

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 9709

&lt;212&gt; DNA

&lt;213&gt; Human immunodeficiency virus

&lt;400&gt; 1

tggaagggtc	aatttgggtc	caaaaaagac	aagagatcct	tgatctgtgg	atctaccaca	60
cacaaggcta	cttcctgat	tggcagaact	acacaccagg	gccagggatc	agatatccac	120
tgacctttgg	atggtgcttc	aagttagtac	cagttgaacc	agagcaagta	gaagaggcca	180
aataaggaga	gaagaacagc	ttgttacacc	ctatgagcca	gcatgggatg	gaggacccgg	240
agggagaagt	attagtgtgg	aagtttgaca	gcctcctagc	atttcgtcac	atggcccag	300
agctgcatcc	ggagtactac	aaagactgct	gacatcgagc	tttctacaag	ggactttccg	360
ctggggactt	tccagggagg	tgtggcctgg	gcgggactgg	ggagtggcga	gccctcagat	420
gctacatata	agcagctgct	ttttgcctgt	actgggtctc	tctgggttaga	ccagatctga	480
gcctgggagc	tctctggcta	actagggaa	ccactgctta	agcctcaata	aagcttgcc	540
tgagtgtcta	aagtagtgtg	tgcccgtctg	ttgtgtgact	ctggtaacta	gagatccctc	600
agaccctttt	agtcagtgtg	gaaaatctct	agcagtggtg	cccgaacagg	gacttgaaag	660
cgaaagtaaa	gccagaggag	atctctcgac	gcaggactcg	gcttgctgaa	gcgcgcacgg	720
caagaggcga	ggggcggcga	ctggtgagta	cgccaaaaat	tttgactagc	ggaggctaga	780
aggagagaga	tgggtgctgag	agcgtcggt	ttaagcgggg	gagaattaga	taaatgggaa	840
aaaattcgg	taaggccagg	gggaaagaaa	caatataaac	taaaacatat	agtatgggca	900
agcagggagc	tagaacgatt	cgcagttaat	cctggccttt	tagagacatc	agaaggctgt	960
agacaaatac	tgggacagct	acaaccatcc	cttcagacag	gatcagaaga	acttagatca	1020
ttatataata	caatagcagt	cctctattgt	gtgcatcaaa	ggatagatgt	aaaagacacc	1080
aaggaagcct	tagataagat	agaggaagag	caaaacaaaa	gtaagaaaaa	ggcacagcaa	1140
gcagcagctg	acacaggaaa	caacagccag	gtcagccaaa	attaccctat	agtgcagaa	1200
ctccaggggc	aaatggtaca	tcaggccata	tcacctagaa	ctttaaatgc	atgggtaaaa	1260
gtagtagaag	agaaggcttt	cagcccgaga	gtaataccca	tgttttcagc	attatcagaa	1320
ggagccaccc	cacaagattt	aaataccatg	ctaaacacag	tggggggaca	tcaagcagcc	1380
atgcaaattg	taaaagagac	catcaatgag	gaagctgcag	aatgggatag	attgcatcca	1440
gtgcatgcag	ggcctattgc	accaggccag	atgagagaac	caaggggaag	tgacatagca	1500
ggaactacta	gtacccttca	ggaacaaata	ggatggatga	cacataatcc	acctatccca	1560
gtaggagaaa	tctataaaa	atggataatc	ctgggattaa	ataaaaatag	aagaatgtat	1620
agccctacca	gcattcttga	cataagacaa	ggaccaaagg	aaccctttag	agactatgta	1680
gaccgattct	ataaaactct	aagagccgag	caagcttcac	aagaggtaaa	aaattggatg	1740
acagaaacct	tgtttggtcca	aaatgcgaac	ccagattgta	agactatttt	aaaagcattg	1800
ggaccaggag	cgacactaga	agaaatgatg	acagcatgtc	agggagtggg	gggacccggc	1860
cataaagcaa	gagtttttgg	tgaagcaatg	agccaagtaa	caaattccag	taccataatg	1920

atacagaaag	gcaatTTTTag	gaaccaaaga	aagactgtta	agtgtttcaa	ttgtggcaaa	1980
gaagggcaca	tagccaaaaa	ttgcagggcc	cctaggaaaa	agggctgttg	gaaatgtgga	2040
aaggaaggac	accaaataag	agattgtact	gagagacagg	ctaatttttt	agggagatc	2100
tggccttccc	acaaggggaag	gccaggggaat	tttcttcaga	gcagaccaga	gccaacagcc	2160
ccaccagaag	agagcttcag	gtttggggaa	gagacaacaa	ctccctctca	gaagcaggag	2220
ccgatagaca	aggaactgta	tccttttagct	tccttcagat	cactcttttg	cagcgacccc	2280
tcgtcacaa	aaagataggg	gggcaattaa	aggaagctct	attagatata	ggagcagatg	2340
atacagtatt	agaagaaatg	aatttgccag	gaagatggaa	acaaaaaatg	atagggggaa	2400
ttggagggtt	tatcaaagta	ggacagtatg	atcagatact	catagaaatc	tgcggacata	2460
aagctatagg	tacagtatta	gtaggaccta	cacctgtcaa	cataattgga	agaaatctgt	2520
tgactcagat	tggtcgcact	ttaaatTTTT	ccattagtcc	tattgagact	gtaccagtaa	2580
aattaaagcc	aggaatggat	ggcccaaaaag	ttaaacaatg	gccattgaca	gaagaaaaaa	2640
taaaagcatt	agtagaaatt	tgtacagaaa	tggaaaagga	agggaaaatt	tcaaaaattg	2700
ggcctgaaaa	tccatacaat	actccagtat	ttgccataaa	gaaaaaagac	agtactaaat	2760
ggagaaaatt	agtagatttc	agagaactta	ataagagaa	tcaagatttc	tgggaagttc	2820
aattaggaat	accacatcct	gcagggttaa	aacagaaaaa	atcagtaaca	gtactggatg	2880
tgggcgatgc	atatttttca	gttcccttag	ataaagactt	caggaagtat	actgcattta	2940
ccatacctag	tataaacaat	gagacaccag	ggattagata	tcagtacaat	gtgcttccac	3000
agggatggaa	aggatcacca	gcaatatccc	agtgtagcat	gacaaaaatc	ttagagcctt	3060
ttagaaaaca	aaatccagac	atagtcctct	atcaatcat	ggatgatttg	tatgtaggat	3120
ctgacttaga	aatagggcag	catagaacaa	aaatagagga	actgagacaa	catctgttga	3180
ggtggggatt	taccacacca	gacaaaaaac	atcagaaaga	acctccattc	ctttggatgg	3240
gttatgaact	ccatcctgat	aaatggacag	tacagcctat	agtgtgtcca	gaaaaggaca	3300
gctggactgt	caatgacata	cagaaattag	tgggaaaatt	gaattgggca	agtcagattt	3360
atgcagggat	taaagtaagg	caattatgta	aacttcttag	gggaaccaa	gcactaacag	3420
aagtagtacc	actaacagaa	gaagcagagc	tagaactggc	agaaaacagg	gagattctaa	3480
aagaaccggt	acatggagtg	tattatgacc	catcaaaaga	cttaatagca	gaaatacaga	3540
agcaggggca	aggccaatgg	acatatcaaa	tttatcaaga	gccattttaa	aatctgaaaa	3600
caggaaaata	tgcaagaatg	aagggtgccc	acactaatga	tgtgaaacaa	ttaacagagg	3660
cagtacaaaa	aatagccaca	gaaagcatag	taatatgggg	aaagactcct	aaattttaa	3720
taccatacaca	aaaggaaaca	tgggaagcat	ggtggacaga	gtattggcaa	gccacctgga	3780
ttcctgagt	ggagtttgtc	aatacccttc	ccttagtgaa	gttatggtac	cagtttagaga	3840
aagaacccat	aataggagca	gaaactttct	atgtagatgg	ggcagccaat	agggaaacta	3900
aattaggaaa	agcaggatat	gtaactgaca	gaggaagaca	aaaagtgtgc	cccctaacgg	3960
acacaacaaa	tcagaagact	gagttacaag	caattcatct	agctttgcag	gattcgggat	4020
tagaagtaaa	catagtgaca	gactcacaat	atgcattggg	aatcattcaa	gcacaaccag	4080
ataagagtga	atcagagtta	gtcagtcaaa	taatagagca	gttaataaaa	aaggaaaaag	4140
tctacctggc	atgggtacca	gcacacaaag	gaattggagg	aatgaacaa	gtagatgggt	4200
tggtcagtgc	tggaatcagg	aaagtactat	ttttagatgg	aatagataag	gccaagaagg	4260
aacatgagaa	atatcacagt	aattggagag	caatggctag	tgatttttaac	ctaccacctg	4320
tagtagcaaa	agaaatagta	gccagctgtg	ataaatgtca	gctaaaaggg	gaagccatgc	4380
atggacaagt	agactgtagc	ccaggaatat	ggcagctaga	ttgtacacat	ttagaaggaa	4440
aagttatctt	ggtagcagtt	catgtagcca	gtggatatat	agaagcagaa	gtaattccag	4500
cagagacagg	gcaagaaaca	gcatacttcc	tcttaaaatt	agcaggaaga	tggccagtaa	4560
aaacagtaca	tacagacaat	ggcagcaatt	tcaccagtac	tacagttaag	gccgcctgtt	4620
ggtgggcggg	gatcaagcag	gaatttggca	ttccctacaa	tcccaaaagt	caaggagtaa	4680
tagaatctat	gaataaagaa	ttaaagaaaa	ttataggaca	ggtaagagat	caggctgaac	4740
atcttaagac	agcagtacaa	atggcagtat	tcattccaca	ttttaaaaga	aaagggggga	4800
ttggggggta	cagtgcaggg	gaaagaatag	tagacataat	agcaacagac	atacaaaacta	4860
aagaattaca	aaaacaaatt	acaaaaattc	aaaattttcg	ggtttattac	agggacagca	4920
gagatccagt	ttggaaagga	ccagcaaagc	tcctctggaa	aggtgaaggg	gcagtagtaa	4980
tacaagataa	tagtgacata	aaagtagtgc	caagaagaaa	agcaaagatc	atcagggtat	5040
atggaaaaca	gatggcaggt	gatgattgtg	tggcaagtag	acaggatgag	gattaacaca	5100
tggaaaagat	tagtaaaaca	ccatatgtat	atttcaagga	aagctaagga	ctgggtttat	5160
agacatcact	atgaaagtac	taatccaaaa	ataagttcag	aagtacacat	cccactaggg	5220
gatgctaaat	tagtaataac	aacatattgg	ggtctgcata	caggagaaag	agactggcat	5280
ttgggtcagg	gagtcctcat	agaatggagg	aaaaagagat	atagcacaca	agtagaccct	5340

gacctagcag	accaactaat	tcatctgcac	tatttttgatt	gtttttcaga	atctgctata	5400
agaaataacca	tattaggacg	tatagttagt	cctaggtgtg	aatatcaagc	aggacataac	5460
aaggtaggat	ctctacagta	cttggcacta	gcagcattaa	taaaaccaa	acagataaag	5520
ccacctttgc	ctagtgttag	gaaactgaca	gaggacagat	ggaacaagcc	ccagaagacc	5580
aagggccaca	gagggagcca	tacaatgaat	ggacactaga	gcttttagag	gaacttaaga	5640
gtgaagctgt	tagacatttt	cctaggatat	ggctccataa	cttaggacaa	catatctatg	5700
aaacttacgg	ggatacttgg	gcaggagtgg	aagccataat	agaattcttg	caacaactgc	5760
tgtttatcca	tttcagaatt	gggtgtcgac	atagcagaat	aggcgttact	cgacagagga	5820
gagcaagaaa	tggagccagt	agatcctaga	ctagagccct	ggaagcatcc	aggaagtcag	5880
cctaaaactg	cttgtacca	ttgctattgt	aaaaagtgtt	gctttcattg	ccaagtttgt	5940
ttcatgacaa	aagccttagg	catctcctat	ggcaggaaga	agcggagaca	gcgacgaaga	6000
gctcatcaga	acagtcagac	tcatcaagct	tctctatcaa	agcagtaagt	agtacatgta	6060
atgcaaccta	taatagtagc	aatagtagca	ttagtagtag	caataataat	agcaatagtt	6120
gtgtggtcca	tagtaatcat	agaatatagg	aaaatatata	gacaaagaaa	aatagacagg	6180
ttaattgata	gactaataga	aagagcagaa	gacagtggca	atgagagtga	aggagaagta	6240
tcagcacttg	tggagatggg	ggtggaaatg	gggcaccatg	ctccttgga	tattgatgat	6300
ctgtagtgt	acagaaaaat	tgtgggtcac	agtctattat	ggggtacctg	tgtggaagga	6360
agcaaccacc	actctatttt	gtgcatcaga	tgctaaagca	tatgatacag	aggtacataa	6420
tgtttggggc	acacatgcct	gtgtaccac	agaccccaac	ccacaagaag	tagtattggg	6480
aaatgtgaca	gaaaatttta	acatgtggaa	aaatgacatg	gtagaacaga	tgcatgagga	6540
tataatcagt	ttatgggatc	aaagcctaaa	gccatgtgta	aaattaaccc	cactctgtgt	6600
tagtttaag	tgcactgatt	tgaagaatga	tactaatacc	aatagtagta	gctgggagaat	6660
gataatggag	aaaggagaga	taaaaaactg	ctctttcaat	atcagcacia	gcataagaga	6720
taaggtgcag	aaagaatatg	cattctttta	taaacttgat	atagtaccaa	tagataatac	6780
cagctatagg	ttgataagtt	gtaacacctc	agtcattaca	caggcctgtc	caaaggatc	6840
ctttgagcca	attcccatac	attattgtgc	cccggctggg	tttgcgattc	taaaatgtaa	6900
taataagacg	ttcaatggaa	caggaccatg	tacaaatgtc	agcacagtac	aatgtacaca	6960
tggaaatcagg	ccagtagtat	caactcaact	gctgttaa	ggcagtctag	cagaagaaga	7020
tgtagtaatt	agatctgcca	atttcacaga	caatgctaaa	accataatag	tacagctgaa	7080
cacatctgta	gaaattaatt	gtacaagacc	caacaacaat	acaagaaaaa	gtatccgtat	7140
ccagagggga	ccagggagag	catttggttac	aataggaaaa	ataggaaata	tgagacaagc	7200
acattgtaac	attagtagag	caaaatggaa	tgccacttta	aaacagatag	ctagcaaatt	7260
aagagaacaa	tttggaata	ataaaacaat	aatctttaag	caatcctcag	gaggggaccc	7320
agaaattgta	acgcacagtt	ttaattgtgg	aggggaattt	ttctactgta	attcaacaca	7380
actgtttaat	agtacttggt	ttaatagtac	ttggagtact	gaaggggtcaa	ataacactga	7440
aggaagtgc	acaatcacac	tcccatgcag	aataaaacaa	tttataaaca	tgtggcagga	7500
agtaggaaaa	gcaatgtatg	ccccctccat	cagtggacaa	attagatggt	catcaaata	7560
tactgggctg	ctattaacaa	gagatggtgg	taataacaac	aatgggtccg	agatcttcag	7620
acctggagga	ggcgatatga	gggacaattg	gagaagtga	ttatataaat	ataaagtagt	7680
aaaaattgaa	ccattaggag	tagcaccac	caaggcaaag	agaagagtgg	tgagagaga	7740
aaaaagagca	gtgggaatag	gagctttggt	ccctgggttc	ttgggagcag	caggaagcac	7800
tatgggctgc	acgtcaatga	cgctgacggt	acaggccaga	caattattgt	ctgatatagt	7860
gcagcagcag	aacaatttgc	tgagggttat	tgaggcgcaa	cagcatctgt	tgcaactcac	7920
agtctggggc	atcaaacagc	tccaggcaag	aatcctggct	gtggaaagat	acctaaagga	7980
tcaacagctc	ctggggattt	gggggttgctc	tggaaaactc	atttgcacca	ctgctgtgcc	8040
ttggaatgct	agttggagta	ataaatctct	ggaacagatt	tggataaaca	tgacctggat	8100
ggagtgggac	agagaaatta	acaattacac	aagcttaata	cactccttaa	ttgaagaatc	8160
gcaaaaccag	caagaaaaga	atgaacaaga	attattggaa	ttagataaat	gggcaagttt	8220
gtggaattgg	tttaacataa	caaattggct	gtggtatata	aaattattca	taatgatagt	8280
aggaggcttg	gtaggtttta	gaatagtttt	tgctgtactt	tctatagtga	atagagttag	8340
gcagggatat	tcaccattat	cgtttcagac	ccacctccca	atcccagagg	gacccgacag	8400
gcccgaagga	atagaagaag	aaggtggaga	gagagacaga	gacagatcca	ttcgattagt	8460
gaacggatcc	ttagcactta	tctgggacga	tctgcggagc	ctgtgcctct	tcagctacca	8520
ccgcttgaga	gacttactct	tgattgtaac	gaggattgtg	gaacttctgg	gacgcagggg	8580
gtgggaagcc	ctcaaatatt	ggtggaatct	cctacagtat	tggagtccag	aactaaagaa	8640
tagtgctgtt	aacttgctca	atgccacagc	catagcagta	gctgagggga	cagatagggg	8700
tatagaagta	ttacaagcag	cttatagagc	tattcgccac	atacctagaa	gaataagaca	8760

```

gggcttggaa aggattttgc tataagatgg gtggcaagtg gtcaaaaagt agtgtgattg 8820
gatggcctgc tgtaagggaa agaatgagac gagctgagcc agcagcagat ggggtgggag 8880
cagtatctcg agacctagaa aaacatggag caatcacaag tagcaataca gcagctaaca 8940
atgctgcttg tgcttggtta gaagcacaag aggaggaaga ggtggggtttt ccagtcacac 9000
ctcaggtacc ttttaagacca atgacttaca aggcagctgt agatcttagc cactttttta 9060
aagaaaaggg gggactggaa gggctaattc actcccaaag aagacaagat atccttgatc 9120
tgtggatcta ccacacacaa ggctacttcc ctgattggca gaactacaca ccagggccag 9180
gggtcagata tccactgacc tttggatggg gctacaagct agtaccagtt gagccagata 9240
aggtagaaga ggccaataaa ggagagaaca ccagcttggt acaccctgtg agcctgcatg 9300
gaatggatga ccctgagaga gaagtgttag agtggagggt tgacagccgc ctagcatttc 9360
atcacgtggc ccgagagctg catccggagt acttcaagaa ctgctgacat cgagcttgct 9420
acaagggact ttccgctggg gactttccag ggaggcgtgg cctgggaggg actggggagt 9480
ggcgagccct cagatgctgc atataagcag ctgctttttg cctgtactgg gtctctcttg 9540
ttagaccaga tctgagcctg ggagctctct ggctaactag ggaaccact gcttaagcct 9600
caataaagct tgccttgagt gcttcaagta gtgtgtgccc gtctgttggt tgactctggg 9660
aactagagat ccctcagacc cttttagtca gtgtggaaaa tctctagca 9709

```

<210> 2  
 <211> 854  
 <212> PRT  
 <213> Human immunodeficiency virus

<220>  
 <223> Envelope Polyprotein

<400> 2  
 Met Arg Val Lys Glu Lys Tyr Gln His Leu Trp Arg Trp Gly Trp Lys  
 1 5 10 15  
 Trp Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Thr Glu  
 20 25 30  
 Lys Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala  
 35 40 45  
 Thr Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu  
 50 55 60  
 Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn  
 65 70 75 80  
 Pro Gln Glu Val Val Leu Val Asn Val Thr Glu Asn Phe Asn Met Trp  
 85 90 95  
 Lys Asn Asp Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp  
 100 105 110  
 Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Ser  
 115 120 125  
 Leu Lys Cys Thr Asp Leu Lys Asn Asp Thr Asn Thr Asn Ser Ser Ser  
 130 135 140  
 Gly Arg Met Ile Met Glu Lys Gly Glu Ile Lys Asn Cys Ser Phe Asn  
 145 150 155 160

Ile	Ser	Thr	Ser	Ile	Arg	Asp	Lys	Val	Gln	Lys	Glu	Tyr	Ala	Phe	Phe	
				165					170					175		
Tyr	Lys	Leu	Asp	Ile	Val	Pro	Ile	Asp	Asn	Thr	Ser	Tyr	Arg	Leu	Ile	
			180					185					190			
Ser	Cys	Asn	Thr	Ser	Val	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	
		195					200					205				
Glu	Pro	Ile	Pro	Ile	His	Tyr	Cys	Ala	Pro	Ala	Gly	Phe	Ala	Ile	Leu	
	210					215					220					
Lys	Cys	Asn	Asn	Lys	Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Thr	Asn	Val	
225					230					235					240	
Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile	Arg	Pro	Val	Val	Ser	Thr	Gln	
				245					250					255		
Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Asp	Val	Val	Ile	Arg	Ser	
			260					265					270			
Ala	Asn	Phe	Thr	Asp	Asn	Ala	Lys	Thr	Ile	Ile	Val	Gln	Leu	Asn	Thr	
		275					280					285				
Ser	Val	Glu	Ile	Asn	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	
	290					295					300					
Ile	Arg	Ile	Gln	Arg	Gly	Pro	Gly	Arg	Ala	Phe	Val	Thr	Ile	Gly	Lys	
305					310					315					320	
Ile	Gly	Asn	Met	Arg	Gln	Ala	His	Cys	Asn	Ile	Ser	Arg	Ala	Lys	Trp	
				325					330					335		
Asn	Ala	Thr	Leu	Lys	Gln	Ile	Ala	Ser	Lys	Leu	Arg	Glu	Gln	Phe	Gly	
			340					345					350			
Asn	Asn	Lys	Thr	Ile	Ile	Phe	Lys	Gln	Ser	Ser	Gly	Gly	Asp	Pro	Glu	
		355					360					365				
Ile	Val	Thr	His	Ser	Phe	Asn	Cys	Gly	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	
	370					375					380					
Ser	Thr	Gln	Leu	Phe	Asn	Ser	Thr	Trp	Phe	Asn	Ser	Thr	Trp	Ser	Thr	
385					390					395					400	
Glu	Gly	Ser	Asn	Asn	Thr	Glu	Gly	Ser	Asp	Thr	Ile	Thr	Leu	Pro	Cys	
			405						410					415		
Arg	Ile	Lys	Gln	Phe	Ile	Asn	Met	Trp	Gln	Glu	Val	Gly	Lys	Ala	Met	
			420					425					430			
Tyr	Ala	Pro	Pro	Ile	Ser	Gly	Gln	Ile	Arg	Cys	Ser	Ser	Asn	Ile	Thr	
	435						440					445				
Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Asn	Asn	Asn	Asn	Gly	Ser	Glu	

450					455					460					
Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asp	Asn	Trp	Arg	Ser	Glu
465					470					475					480
Leu	Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile	Glu	Pro	Leu	Gly	Val	Ala	Pro
				485					490					495	
Thr	Lys	Ala	Lys	Arg	Arg	Val	Val	Gln	Arg	Glu	Lys	Arg	Ala	Val	Gly
			500					505					510		
Ile	Gly	Ala	Leu	Phe	Leu	Gly	Phe	Leu	Gly	Ala	Ala	Gly	Ser	Thr	Met
		515					520					525			
Gly	Cys	Thr	Ser	Met	Thr	Leu	Thr	Val	Gln	Ala	Arg	Gln	Leu	Leu	Ser
	530					535					540				
Asp	Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	Gln
545					550					555					560
Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	Ala
				565					570						575
Arg	Ile	Leu	Ala	Val	Glu	Arg	Tyr	Leu	Lys	Asp	Gln	Gln	Leu	Leu	Gly
			580					585					590		
Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ala	Val	Pro	Trp
		595					600					605			
Asn	Ala	Ser	Trp	Ser	Asn	Lys	Ser	Leu	Glu	Gln	Ile	Trp	Asn	Asn	Met
	610					615					620				
Thr	Trp	Met	Glu	Trp	Asp	Arg	Glu	Ile	Asn	Asn	Tyr	Thr	Ser	Leu	Ile
625					630					635					640
His	Ser	Leu	Ile	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Lys	Asn	Glu	Gln
				645					650					655	
Glu	Leu	Leu	Glu	Leu	Asp	Lys	Trp	Ala	Ser	Leu	Trp	Asn	Trp	Phe	Asn
			660					665					670		
Ile	Thr	Asn	Trp	Leu	Trp	Tyr	Ile	Lys	Leu	Phe	Ile	Met	Ile	Val	Gly
		675					680					685			
Gly	Leu	Val	Gly	Leu	Arg	Ile	Val	Phe	Ala	Val	Leu	Ser	Ile	Val	Asn
	690					695					700				
Arg	Val	Arg	Gln	Gly	Tyr	Ser	Pro	Leu	Ser	Phe	Gln	Thr	His	Leu	Pro
705					710					715					720
Ile	Pro	Arg	Gly	Pro	Asp	Arg	Pro	Glu	Gly	Ile	Glu	Glu	Glu	Gly	Gly
				725					730					735	
Glu	Arg	Asp	Arg	Asp	Arg	Ser	Ile	Arg	Leu	Val	Asn	Gly	Ser	Leu	Ala
			740					745					750		

Leu Ile Trp Asp Asp Leu Arg Ser Leu Cys Leu Phe Ser Tyr His Arg  
755 760 765

Leu Arg Asp Leu Leu Leu Ile Val Thr Arg Ile Val Glu Leu Leu Gly  
770 775 780

Arg Arg Gly Trp Glu Ala Leu Lys Tyr Trp Trp Asn Leu Leu Gln Tyr  
785 790 795 800

Trp Ser Gln Glu Leu Lys Asn Ser Ala Val Asn Leu Leu Asn Ala Thr  
805 810 815

Ala Ile Ala Val Ala Glu Gly Thr Asp Arg Val Ile Glu Val Leu Gln  
820 825 830

Ala Ala Tyr Arg Ala Ile Arg His Ile Pro Arg Arg Ile Arg Gln Gly  
835 840 845

Leu Glu Arg Ile Leu Leu  
850

<210> 3

<211> 107

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:  
oligonucleotide for cloning

<400> 3

aagatgtagt aattagatct gccaatattca cagacaatgc taaaaccata atagtacagc 60  
tgaacacatc gttagaaatt aattgtacaa gaccaacaa caataca 107

<210> 4

<211> 120

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence:  
oligonucleotide for cloning

<220>

<221> misc\_feature

<222> (97)..(99)

<223> Sequence at this position: (GA)(AT)(GATC), ie.  
base at position 97 can be G or A, base at  
position 98 can be A or T, and base at  
position 99 can be G, A, T or C.

<400> 4

ttttgctcta gaaatgttac aatgtgcttg tcttatgtct cctgttgcag cttctgttgc 60  
atgaaatgct ctccctgggc cgatatggat actatgrwnt tttcttgtat tgttgttggg 120

<210> 5  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 5  
ccatgtacaa atgtcag 17

<210> 6  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 6  
aaaactgtgc gttacaa 17

<210> 7  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 7  
gtaaaacgac ggccagt 17

<210> 8  
<211> 17  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence:  
sequencing primer

<400> 8  
caggaaacag ctatgac 17



<210> 9  
<211> 2148  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence: synthetic DNA

<220>  
<221> misc\_feature  
<222> (3)..(9)  
<223> BstEII cleavage site

<220>  
<221> misc\_feature  
<222> (2143)..(2148)  
<223> BamHI cleavage site

<400> 9  
tgggtcaccg tctattatgg ggtgcctgtg tggaaggaag caaccaccac tctatthttgt 60  
gcatcagatg ctaaagcata tgatacagag gtacataatg tttggggccac acatgcctgt 120  
gtaccacag accccaaccc acaagaagta gtatttgtaa atgtgacaga aaatttttaac 180  
atgtggaaaa atgacatggt agaacagatg catgaggata taatcagttt atgggatcaa 240  
agccttaagc catgtgtaaa attaacccca ctctgtgtta gtttaaagtg cactgatttg 300  
aagaatgata ctaataccaa tagtagtagc gggagaatga taatggagaa aggagagata 360  
aaaaactgca gcttcaatat cagcacaagc ataagagata aggtgcagaa agaatatgca 420  
ttctttttata aacttgatat agtaccaata gataatacca gctatagggt gataagttgt 480  
aacacctcag tgatcacaca ggcctgtcca aagggtatcct ttgagccaat tcccatacat 540  
tattgtgccc cggctgggtt tgcgattcta aaatgtaata ataagacgtt caatggaaca 600  
ggaccatgta caaatgtcag cacagtacaa tgtacacatg gaattcgacc agtagtatca 660  
actcaactgc tgttaaatgg cagtctagca gaagaagatg tagtaattag atctgccaat 720  
ttcacagaca atgctaaaac cataatagta cagctgaaca catctgtaga aattaattgt 780  
acaagaccca acaacaatac aagaaaaagt atccgtatcc agaggggacc agggagagca 840  
tttgttacaa taggaaaaat aggaaatatg agacaagcac attgtaacat ttctagagca 900  
aaatggaatg ccacttttaa acagatagct agcaaattaa gagaacaatt tggaaataat 960  
aaaacaataa tctttaagca gtcattccga ggggaccag aaattgtaac gcacagtttt 1020  
aattgtggag ggaatttttt ctactgtaat tcaacacaac tgtttaatag tacttggttt 1080  
aatagtactt ggagtactga aggggtcaat aacactgaag gaagtgcac aatcacactc 1140  
ccatgcagaa taaaacaatt tataaacatg tggcaggaag taggaaaagc aatgtatgcc 1200  
cctcccatca gtggccaaat tagatgttca tcaaatatta ctgggctgct attaaactcg 1260  
gatgggtgta ataacaacaa tgggtccgag attttcagac ctggaggagg cgatatgagg 1320  
gataattgga gaagtgaatt atataaatat aaagtagtaa aaattgaacc attaggagta 1380  
gcaccacca aggcaaagag acgcgtggtg cagagagaaa agcgcgcagt gggaatagga 1440  
gctctgttcc ttgggttctt gggagcagca ggaagcacta tgggcgcagc gtcaatgacg 1500  
ctgacggtac aggcagaca attattgtct gatatagtgc agcagcagaa caatttgctg 1560  
agggcaattg aggcgaaca gcatctgttg caactcacag tctggggcat caaacagctc 1620  
caggcaagaa tcctggctgt ggaaagatac ctaaaggatc aacagctcct ggggatttgg 1680  
ggttgctctg gaaaactcat ttgcaccact gctgtgcctt ggaatgctag ttggagtaat 1740  
aaatctctgg aacagatttg gaataacatg acctggatgg agtgggacag agaaattaac 1800  
aattacacaa gcttaataca ctcttaatt gaagaatcgc aaaaccagca agaaaagaat 1860  
gaacaagaat tattggaatt agataaatgg gcaagtttgt ggaattggtt taacataaca 1920  
aattggctgt ggtatataaa attattcata atgatagtag gaggcttggg aggtttaaga 1980  
atagtttttg ctgtactttc tatagtgaat agagttaggc agggatatcc accattatcg 2040  
tttcagaccc acctcccaat cccgagggga cccgacaggc ccgaaggaat agaagaagaa 2100  
ggtggagaga gagacagaga cagatccatt cgattagtga acggatcc 2148

<210> 10  
<211> 6229  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of the artificial sequence: synthetic DNA

<220>  
<221> sig\_peptide  
<222> (1293)..(1295)  
<223> env ATG

<220>  
<221> misc\_feature  
<222> (1377)..(1379)  
<223> env AGT, gp120 start

<220>  
<221> misc\_feature  
<222> (1397)..(1403)  
<223> BstEII cleavage site

<220>  
<221> misc\_feature  
<222> (3537)..(3542)  
<223> BamHI cleavage site

<220>  
<221> misc\_feature  
<222> (3855)..(3857)  
<223> env TAA, stop

<400> 10  
ctgacgcgcc ctgtagcggc gcattaagcg cggcgggtgt ggtgggttacg cgcagcgtga 60  
ccgctacact tgccagcgcc ctagcgcccc ctcctttcgc tttcttcctt tcctttctcg 120  
ccacgttcgc cggctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat 180  
ttagtgcttt acggcacctc gaccccaaaa aacttgatta gggatgatgg tcacgtagtg 240  
ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt ggagtccacg ttctttaata 300  
gtggactctt gttccaaact ggaacaacac tcaaccctat ctcggtctat tcttttgatt 360  
tataagggat tttgccgatt tcggcctatt ggtaaataaa tgagctgatt taacaaaaat 420  
  
ttaacgcgaa ttttaacaaa atattaacgc ttacaatttc cattcgccat tcaggetgcg 480  
caactgttgg gaagggcgat cgggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 540  
gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg 600  
taaaacgacg gccagtgagc gtctagttat taatagtaat caattacggg gtcattagtt 660  
catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc gcctggctga 720  
ccgccaacg acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca 780  
atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttggca 840  
  
gtacatcaag tgtatcatat gccaaagtacg ccccctattg acgtcaatga cggtaaattg 900  
cccgcctggc attatgcca gtacatgacc ttatgggact ttctacttg gcagtacatc 960  
tacgtattag tcatcgctat taccatgggt atgcggtttt ggcagtacat caatgggagt 1020  
ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt 1080  
ttgttttggc accaaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg 1140

acgcaaattg	gcggtaggcg	tgtacgggtg	gaggtctata	taagcagagc	tcgttttagtg	1200
aaccgtcaga	tcgcctggag	acgccatcca	cgctgttttg	acctccatag	aagacaccgg	1260
gacaattcga	gctcggtagc	gtcgacgcca	ccatgagagt	gaaggagaag	tatcagcact	1320
tgtggagatg	gggggtgaaa	tggggcacca	tgctccttgg	gataattgatg	atctgtagtg	1380
ctacagaaaa	attgtgggtc	accgtctatt	atggggtagc	tgtgtggaag	gaagcaacca	1440
ccactctatt	ttgtgcatca	gatgctaaag	catatgatac	agaggtacat	aatgtttggg	1500
ccacacatgc	ctgtgtaccc	acagacccca	accacacaaga	agtagtattg	gtaaatgtga	1560
cagaaaaatt	taacatgtgg	aaaaatgaca	tggtagaaca	gatgcatgag	gatataatca	1620
gtttatggga	tcaaagccta	aagccatgtg	taaaattaac	cccactctgt	gttagtttaa	1680
agtgcactga	tttgaagaat	gatactaata	ccaatagtag	tagcgggaga	atgataatgg	1740
agaaaggaga	gataaaaaac	tgctctttca	atatcagcac	aagcataaga	gataagggtc	1800
agaaagaata	tgcattcttt	tataaacttg	atatagtagc	aatagataat	accagctata	1860
ggttgataag	ttgtaacacc	tcagtcatta	cacaggcctg	tccaaaggta	tcctttgagc	1920
caattcccat	acattattgt	gccccggctg	gttttgcat	tctaaaatgt	aataataaga	1980
cgttcaatgg	aacaggacca	tgtacaaatg	tcagcacagt	acaatgtaca	catggaatca	2040
ggccagtagt	atcaactcaa	ctgctgttaa	atggcagctc	agcagaagaa	gatgtagtaa	2100
ttagatctgc	caatttcaca	gacaatgcta	aaaccataat	agtacagctg	aacacatctg	2160
tagaaattaa	ttgtacaaga	cccaacaaca	atacaagaaa	aagtatccgt	atccagaggg	2220
gaccagggag	agcatttgtt	acaataggaa	aaataggaaa	tatgagacaa	gcacattgtg	2280
acattagtag	agcaaaatgg	aatgccactt	taaaacagat	agctagcaaa	ttaagagaac	2340
aatttggtaaa	taataaaaaca	ataatcttta	agcaatcctc	aggaggggac	ccagaaattg	2400
taacgcacag	ttttaattgt	ggaggggaat	ttttctactg	taattcaaca	caactgttta	2460
atagtacttg	gtttaatagt	acttggaagt	ctgaagggtc	aaataacact	gaagggaagt	2520
acacaatcac	actcccatgc	agaataaaac	aattttataaa	catgtggcag	gaagtaggaa	2580
aagcaatgta	tgccccctcc	atcagtggac	aaattagatg	ttcatcaaat	attactgggc	2640
tgctattaac	aagagatggg	ggtaataaca	acaatgggtc	cgagatcttc	agacctggag	2700
gaggcgatat	gaggggacaat	tggagaagtg	aattatataa	atataaagta	gtaaaaattg	2760
aaccattagg	agtagcaccc	accaaggcaa	agagaagagt	ggtgcagaga	gaaaaaagag	2820
cagtgggaat	aggagctttg	ttccttgggt	tcttgggagc	agcaggaagc	actatgggct	2880
gcacgtcaat	gacgctgacg	gtacaggcca	gacaattatt	gtctgatata	gtgcagcagc	2940
agaacaattt	gctgagggct	attgaggcgc	aacagcatct	gttgcaactc	acagtctggg	3000
gcatcaaaca	gctccaggca	agaatcctgg	ctgtggaaag	atacctaaag	gatcaacagc	3060
tcctggggat	ttgggggttg	tctggaaaac	tcatttgcac	cactgctgtg	ccttggaatg	3120
ctagttaggag	taataaatct	ctggaacaga	tttggaataa	catgacctgg	atggagtggg	3180
acagagaaat	taacaattac	acaagcttaa	tacactcctt	aattgaagaa	tcgcaaaacc	3240
agcaagaaaa	gaatgaacaa	gaattatttg	aattagataa	atgggcaagt	ttgtggaatt	3300
ggtttaacat	aacaaattgg	ctgtggtata	taaaattatt	cataatgata	gtaggaggct	3360
tggtaggttt	aagaatagtt	tttgctgtac	tttctatagt	gaatagagtt	aggcagggat	3420
attcaccatt	atcgtttcag	accacacctc	caatcccag	gggacccgac	aggcccgag	3480
gaatagaaga	agaaggtgga	gagagagaca	gagacagatc	cattcgatta	gtgaacggat	3540
ccttagcact	tatctgggac	gatctgcgga	gcctgtgcct	cttcagctac	caccgcttga	3600
gagacttact	cttgattgta	acgaggattg	tggaaacttct	gggacgcagg	gggtgggaag	3660
ccctcaaata	ttggtggaat	ctcctacagt	attggagtca	ggaactaaag	aatagtgtct	3720
ttaacttgct	caatgccaca	gccatagcag	tagctgaggg	gacagatagg	gttatagaag	3780
tattacaagc	agcttataga	gctattcgcc	acatacctag	aagaataaga	cagggtcttg	3840
aaaggatttt	gctataagat	gggtggcaag	tggtcaaaaa	gtagtgtgat	tggatggcct	3900
gctgtaaggg	aaagaatgag	acgagctgag	ccagcagcag	atgggggtgg	agcagtatct	3960
cgagatctag	actagaacta	gcttcgatcc	agacatgata	agatacattg	atgagtttgg	4020
acaaaccaca	actagaatgc	agtgaaaaaa	atgctttatt	tgtgaaattt	gtgatgctat	4080
tgctttattt	gtaaccatta	taagctgcaa	taaacaagtt	aacaacaaca	attgcattca	4140
ttttatgttt	cagggttcagg	gggaggtgtg	ggaggttttt	taaagcaagt	aaaacctcta	4200
caaattgtgg	atggctgatt	atgatcctgc	ctcgcgcgtt	tcggtgatga	cggtgaaaac	4260
ctctgacaca	tgcagctccc	ggagacgggc	acagcttgct	tgtaagcgga	tgccgggagc	4320
agacaagccc	gtcagggcgc	gtcagcgggt	gttgccgggt	gtcggggcgc	agccatgacc	4380
cagtcacgta	gcgatagcgg	agtgtatact	ggcttaacta	tgccggcatca	gagcagattg	4440
tactgagagt	gcaccatatg	tcgggcccgc	ttgctggcgt	ttttccatag	gctccgcccc	4500
cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaaaccc	gacaggacta	4560

taaagataacc	aggcggtttcc	ccctggaagc	tccctcgtgc	gctctcctgt	tccgaccctg	4620
ccgcttaccg	gatacctgtc	cgcttttctc	ccttcgggaa	gcgtggcgct	ttctcatagc	4680
tcacgctgta	ggtatctcag	ttcgggtgtag	gtcgttcgct	ccaagctggg	ctgtgtgcac	4740
gaaccccccg	ttcagcccga	ccgctgcgcc	ttatccggta	actatcgtct	tgagtccaac	4800
ccggtaagac	acgacttatc	gccactggca	gcagccactg	gtaacaggat	tagcagagcg	4860
aggatatgtag	gcggtgctac	agagttcttg	aagtgggtggc	ctaactacgg	ctacactaga	4920
aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	aagagttggt	4980
agctcttgat	ccggcaaaaa	aaccaccgct	ggtagcgggtg	gtttttttgt	ttgcaagcag	5040
cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	tacgggggtct	5100
gacgctcagt	ggaacgaaaa	ctcacgttaa	gggatttttg	tcatgagatt	atcaaaaagg	5160
atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	aagtatatat	5220
gagtaaactt	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	ctcagcgatc	5280
tgtctatttc	gttcatccat	agttgcctga	ctccccgtcg	tgtagataac	tacgatacgg	5340
gagggccttac	catctggccc	cagtgtctgca	atgataaccgc	gagaccacg	ctcaccgggt	5400

ccagattttat	cagcaataaa	ccagccagcc	ggaagggccg	agcgcagaag	tggtcctgca	5460
actttatccg	cctccatcca	gtctattaat	tggtgccggg	aagctagagt	aagtagttcg	5520
ccagttaata	gtttgcgcaa	cgttgttgcc	attgctacag	gcatcgtggt	gtcacgctcg	5580
tcgtttggtg	tggtcttcatt	cagctccggt	tcccaacgat	caaggcgagt	tacatgatcc	5640
cccattgtgt	gcaaaaaagc	ggttagctcc	ttcggtcctc	cgatcgttgt	cagaagtaag	5700
ttggccgcag	tgttatcact	catggttatg	gcagcactgc	ataattctct	tactgtcatg	5760
ccatccgtaa	gatgcttttc	tgtgactggt	gagtactcaa	ccaagtcatt	ctgagaatag	5820
tgtatgcggc	gaccgagttg	ctcttgccc	gcgtcaatac	gggataatac	cgcgccacat	5880
agcagaactt	taaaagtgtc	catcattgga	aaacgttctt	cggggcgaaa	actctcaagg	5940
atcttaccgc	tgttgagatc	cagttcgatg	taaccctctc	gtgcacccaa	ctgatcttca	6000
gcatctttta	ctttcaccag	cgtttctggg	tgagcaaaaa	caggaaggca	aaatgccgca	6060
aaaaagggaa	taagggcgac	acggaaatgt	tgaatactca	tactcttcct	ttttcaatat	6120
tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatttga	atgtatttag	6180
aaaaataaac	aaataggggt	tccgcgcaca	tttccccgaa	aagtgccac		6229

<210> 11  
 <211> 860  
 <212> DNA  
 <213> Human immunodeficiency virus

<220>  
 <221> misc\_feature  
 <222> (1)..(860)  
 <223> PI-932 original sequence V1-V2-V3-loop

<400> 11  
 tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60

aacatgtgga	aaaatgacat	ggtagaacag	atgcatgagg	atataatcaa	tttatgggat	120
caaagcctaa	agccatgtgt	aaaactaacc	ccactctgtg	ttacttttaa	ttgactgat	180
gctgatttaa	attgcaataa	tactgattta	aattgcacta	aagctaattt	ggggaaaaat	240
actcataaca	atactattag	tgggaaaata	atagagaaag	tagaaataaa	aaactgctct	300
ttcaagggtca	ccacaggcat	aagggataag	atgcaaaaag	aatatgcact	tttgaataaa	360
cttgatatag	taccaataga	taatgataag	aataatacta	actttatatt	gataagttgt	420
aacacctcga	ccattacaca	ggcctgtcca	aaggatcctt	ttgagccaat	tcccatacat	480
ttttgtgcc	cggctgggtt	tgcgattcta	aagtgtaatg	aaaagagtta	cagtggaaaa	540
ggaccatgta	aaaatgtcag	cacagtacaa	tgtacacatg	gaattaggcc	agtagtgtca	600
actcaactgc	tgttgaatgg	cagtctagca	gaaaaagaag	tagtaattag	atctgagaat	660
ttcacagaca	atgctaaaac	cataatagta	cagctgaagg	aatctgtaaa	cattacttgt	720
ataagacccc	acaacactgt	aacagacagg	atacatatag	ggccagggag	atcatttcat	780
acaacaagaa	aaataaaaag	agatataaga	caagcacatt	gtagccttag	gagaaaagat	840

tggaataaca ctttacaaga

860

<210> 12

<211> 870

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of the artificial sequence: PI-932  
gene cassette, comprising the cleavage sites for  
restriction enzymes BspT1, PstI, BclI, EcoRI,  
BglII, PvuII, XbaII, NheI

<400> 12

```
tgtgtaccca cagaccccaa cccacaaaag gtagtattgg aaaatgtgac agaaaatttt 60
aacatgtgga aaaatgacat ggtagaacag atgcatgagg atataatcaa tttatgggat 120
caaagcctta agccatgtgt aaaactaacc ccactctgtg ttactttaaa ttgcactgat 180
gctgatttaa attgcaataa tactgattta aattgcacta aagctaattt ggggaaaaat 240
actcataact gcagtattag tgggaaaata atagagaaag tagaaataaa aaactgctct 300
ttcaagggtc ccacaggcat aagggataag atgcaaaaag aatatgcact tttgaataaa 360
cttgatatag taccaataga taatgataag aataatacta actttatatt gataagttgt 420
aacacctcgg tgatcacaca ggccctgtcca aagggtatcct ttgagccaat tcccatacat 480
ttttgtgccc cggtctggtt tgcgattcta aagtgtaatg aaaagagtta cagtggaaaa 540
ggaccatgta aaaatgtcag cacagtacaa tgtacacatg gaattcggcc agtagtgtca 600
actcaactgc tgttgaatgg cagtctagca gaaaaagaag tagtaattag atctgagaat 660
ttcacagaca atgctaaaac cataatagta cagctgaagg aatctgtaaa cattacttgt 720
ataagacccc acaacactgt aacagacagg atacatatag ggccagggag atcatttcat 780
acaacaagaa aaataaaagg agatataaga caagcacatt gtagcctttc tagaaaagat 840
tggaataaca ctttacaaga gatagctagc 870
```

<210> 13

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: V3 loop  
sequence of HIV-1 patient isolate

<400> 13

```
Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro
  1             5             10             15

Gly Arg Ala Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln
      20             25             30

Ala His Cys
      35
```

<210> 14

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-903 isolate

<400> 14

Cys Thr Arg Pro Asn Asn Asn Thr Arg Gly Ser Ile His Ile Gly Pro  
1 5 10 15  
Gly Ser Thr Asn Tyr Ala Thr Gly Ala Ile Ile Gly Asp Ile Ser Gln  
20 25 30  
Ala His Cys  
35

<210> 15

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-951 isolate

<400> 15

Cys Thr Arg Pro His Asn Asn Thr Arg Lys Ser Ile Asn Ile Gly Pro  
1 5 10 15  
Gly Arg Ala Trp Tyr Thr Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
20 25 30  
Ala His Cys  
35

<210> 16

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-918 isolate

<400> 16

Cys Thr Arg Pro Ser Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro  
1 5 10 15  
Gly Arg Ala Phe Tyr Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln  
20 25 30  
Ala His Cys  
35

<210> 17

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-970 ISOLATE

<400> 17

Cys Thr Arg Pro Ser Asn Asn Thr Arg Lys Ser Ile His Ile Gly Pro  
1 5 10 15

Gly Arg Ala Phe Tyr Ala Thr Gly Glu Ile Ile Gly Asp Ile Arg Gln  
20 25 30

Ala His Cys  
35

<210> 18

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-990 ISOLATE

<400> 18

Cys Thr Arg Pro Asn Asn Asn Thr Arg Arg Ser Ile Pro Ile Gly Pro  
1 5 10 15

Gly Arg Ala Phe Tyr Thr Thr Gly Asp Ile Val Gly Asp Ile Arg Gln  
20 25 30

Ala His Cys  
35

<210> 19

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-991 ISOLATE

<400> 19

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Pro Ile Ala Pro  
1 5 10 15

Gly Arg Ala Phe Tyr Ala Thr Gly Glu Ile Ile Gly Asn Ile Arg Gln  
20 25 30

Ala His Cys  
35

<210> 20

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-952 ISOLATE

<400> 20

Cys Ile Arg Pro Asn Asn Asn Thr Arg Lys Arg Ile Thr Leu Gly Pro  
1 5 10 15

Gly Arg Val Leu Tyr Thr Thr Gly Glu Ile Ile Gly Asp Ile Arg Lys  
20 25 30

Ala His Cys  
35

<210> 21

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-93 ISOLATE

<400> 21

Cys Ile Arg Pro His Asn Thr Val Thr Asp Arg Ile His Ile Gly Pro  
1 5 10 15

Gly Arg Ser Phe His Thr Thr Arg Lys Ile Lys Gly Asp Ile Arg Gln  
20 25 30

Ala His Cys  
35

<210> 22

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-910 ISOLATE

<400> 22

Cys Thr Arg Pro Ser Ile Gln Lys Arg Arg Ser Val Arg Ile Gly Pro  
1 5 10 15

Gly Arg Ser Phe Ile Ala Ala Arg Ala Ala Thr Gly Asp Ile Arg Lys  
20 25 30

Ala Gln Cys  
35

<210> 23

<211> 35

<212> PRT



<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-911 ISOLATE

<400> 23

Cys Thr Arg Pro Ser Ile Gln Lys Arg Arg Ser Val Arg Ile Gly Pro  
1 5 10 15  
Gly Arg Ser Phe Ile Ala Thr Arg Ala Ala Thr Gly Asp Ile Arg Lys  
20 25 30  
Ala Gln Cys  
35

<210> 24

<211> 35

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: PI-930 ISOLATE

<400> 24

Cys Thr Arg Pro Tyr Arg Asn Ala Lys His Arg Ile Met His Ile Gly  
1 5 10 15  
Pro Gly Arg Ala Phe Tyr Ala Thr Asn Val Lys Gly Asn Ile Lys Gln  
20 25 30  
Ala His Cys  
35

<210> 25

<211> 51

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: NL4-3 V2-Loop

<400> 25

Gly Glu Ile Lys Asn Cys Ser Phe Asn Ile Ser Thr Ser Ile Arg Asp  
1 5 10 15  
Lys Val Gln Lys Glu Tyr Ala Phe Phe Tyr Lys Leu Asp Ile Val Pro  
20 25 30  
Ile Asp Asn Thr Ser Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile  
35 40 45  
Thr Gln Ala  
50

<210> 26  
<211> 59  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: NL4-3 V3-Loop

<400> 26  
Val Gln Leu Asn Thr Ser Val Glu Ile Asn Cys Thr Arg Pro Asn Asn  
1 5 10 15  
Asn Thr Arg Lys Ser Ile Arg Ile Gln Arg Gly Pro Gly Arg Ala Phe  
20 25 30

Val Thr Ile Gly Lys Ile Gly Asn Met Arg Gln Ala His Cys Asn Ile  
35 40 45  
Ser Arg Ala Lys Trp Asn Ala Thr Leu Lys Gln  
50 55

<210> 27  
<211> 37  
<212> PRT  
<213> HIV-1

<220>  
<221> sequence variation in position 1  
<222> (1)..(1)  
<223> Xaa can be N, T, H, D, K, S, E, I, Y, or can be absent

<220>  
<221> sequence variation in position 3  
<222> (3)..(3)  
<223> Xaa can be T, I, V, S, E, L, A, or M

<220>  
<221> sequence variation in position 4

<222> (4)..(4)

<223> Xaa can be R, S, T, or V

<220>

<221> sequence variation in position 5

<222> (5)..(5)

<223> Xaa can be P, L, or T

<220>

<221> sequence variation in position 6

<222> (6)..(6)

<223> Xaa can be N, S, G, H, T, Y, D, or F

<220>

<221> sequence variation in position 7

<222> (7)..(7)

<223> Xaa can be N, T, K, D, I, E, R, H, or Y

<220>

<221> sequence variation in position 8

<222> (8)..(8)

<223> Xaa can be N, Y, K, T, I, H, S, E, F, Q, or can be absent

<220>

<221> sequence variation in position 9

<222> (9)..(9)

<223> Xaa can be T, K, I, R, V, A, or S

<220>

<221> sequence variation in position 10  
<222> (10)..(10)  
<223> Xaa can be R, S, I, K, M, G, N, Q, A, L, T, or E

<220>

<221> sequence variation in position 11  
<222> (11)..(11)  
<223> Xaa can be K, R, Q, S, T, N, E, L, G, or I

<220>

<221> sequence variation in position 12  
<222> (12)..(12)  
<223> Xaa can be S, G, R, H, D, C, K, T, V, or N

<220>

<221> sequence variation in position 13  
<222> (13)..(13)  
<223> Xaa can be I, V, L, M, K, R, T, F, or Y

<220>

<221> sequence variation in position 14  
<222> (14)..(14)  
<223> Xaa can be H, P, N, T, S, Y, R, Q, A, G, I, F, or K

<220>

<221> sequence variation in position 15  
<222> (15)..(15)  
<223> Xaa can be I, L, M, V, F, T, S, K, R, or Y

<220>

<221> sequence variation in position 16

<222> (16)..(16)

<223> Xaa can be G, A, E, R, Q, or T

<220>

<221> sequence variation in position 17

<222> (17)..(17)

<223> Xaa can be P, W, L, Q, F, G, S, A, M, V, R, or T

<220>

<221> sequence variation in position 18

<222> (18)..(18)

<223> Xaa can be G, R, K, E, or W

<220>

<221> sequence variation in position 19

<222> (19)..(19)

<223> Xaa can be R, Q, K, S, G, A, H, E, N, or can be absent

<220>

<221> sequence variation in position 20

<222> (20)..(20)

<223> Xaa can be A, V, T, S, R, G, K, Q, L, N, Y, or can be absent

<220>

<221> sequence variation in position 21

<222> (21)..(21)

<223> Xaa can be F, W, L, V, I, Y, S, M, C, N, or P

<220>

<221> sequence variation in position 22

<222> (22)..(22)

<223> Xaa can be Y, F, H, V, R, I, N, W, D, T, C, L, or S

<220>

<221> sequence variation in position 23

<222> (23)..(23)

<223> Xaa can be T, A, G, R, V, S, K, or can be absent

<220>

<221> sequence variation in position 24

<222> (24)..(24)

<223> Xaa can be T, A, R, I, S, P, N, K, Y, G, or can be absent

<220>

<221> sequence variation in position 25

<222> (25)..(25)

<223> Xaa can be G, E, R, D, K, A, T, Q, S, or can be absent

<220>

<221> sequence variation in position 26

<222> (26)..(26)

<223> Xaa can be E, D, Q, R, G, A, K, N, H, T, S, P, I, M, Y,  
V, or can be absent

<220>

<221> sequence variation in position 27

<222> (27)..(27)

<223> Xaa can be I, V, T, L, K, H, N, R, or can be absent

<220>

<221> sequence variation in position 28

<222> (28)..(28)

<223> Xaa can be I, V, T, A, M, K, L, F, N, Q, or can be absent

<220>

<221> sequence variation in position 29

<222> (29)..(29)

<223> Xaa can be G, E, V, R, or can be absent

<220>

<221> sequence variation in position 30

<222> (30)..(30)

<223> Xaa can be D, N, E, K, G, Q, T, L, H, I, A, V, Y, or can  
be absent

<220>

<221> sequence variation in position 31

<222> (31)..(31)

<223> Xaa can be I, T, V, L, M, K, R, or can be absent

<220>

<221> sequence variation in position 32

<222> (32)..(32)

<223> Xaa can be R, K, S, G, T, I, or can be absent

<220>

<221> sequence variation in position 33

<222> (33)..(33)

<223> Xaa can be Q, K, R, E, L, H, T, or A

<220>

<221> sequence variation in position 34

<222> (34)..(34)

<223> Xaa can be A, P, S, or T

<220>

<221> sequence variation in position 35

<222> (35)..(35)

<223> Xaa can be H, Y, Q, F, R, N, or A

<220>

<221> sequence variation in position 36

<222> (36)..(36)

<223> Xaa can be C, S, K, or can be absent

<220>

<221> sequence variation in position 37

<222> (37)..(37)

<223> Xaa can be N, T, I, D, S, E, K, A, V, Y, H, or can be absent

<400> 27

Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa



1

5

10

15

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
20 25 30

Xaa Xaa Xaa Xaa Xaa  
35